PCT/AU97/00874

Patent Office
Canberra

I, MARIO PERUSSICH, ASSISTANT DIRECTOR PATENT SERVICES, hereby certify that the annexed are true copies of the Provisional specification and drawing(s) as filed on 20 December 1996 in connection with Application No. PO 4275 for a patent by COOPERATIVE RESEARCH CENTRE FOR TROPICAL PLANT PATHOLOGY.

I further certify that the annexed documents are not, as yet, open to public inspection.

PRIORITY DOCUMENT

WITNESS my hand this Fifth
day of January 1998

MARIO PERUSSICH
ASSISTANT DIRECTOR PATENT SERVICES

COOPERATIVE RESEARCH CENTRE FOR TROPICAL PLANT PATHOLOGY

AUSTRALIA

Patents Act 1990

PROVISIONAL SPECIFICATION

Invention Title: Anti-microbial Protein

This invention is described in the following statement:

TECHNICAL FIELD

This invention relates to isolated proteins which exert inhibitory activity on the growth of fungi and bacteria, which fungi and bacteria include some microbial pathogens of plants and animals. The invention also relates to recombinant genes which include sequences encoding the proteins, the expression products of which recombinant genes can contribute to plant cells or cells of other organism's defence against invasion by microbial pathogens. The invention further relates to the use of the proteins and/or genes encoding the proteins for the control of microbes in human and veterinary clinical conditions.

BACKGROUND ART

Microbial diseases of plants are a significant problem to the agricultural and horticultural industries. Plant diseases in general annually cause millions of tonnes of crop losses with fungal and bacterial diseases responsible for significant portions of these losses. One possible way of combating fungal and bacterial diseases is to provide transgenic plants capable of expressing a protein or proteins which in some way increase the resistance of the plant to pathogen attack. A simple strategy is to first identify a protein with anti-microbial activity *in vitro*, to clone or synthesise the DNA sequence encoding the protein, to make a chimaeric gene construct for efficient expression of the protein in plants, to transfer this gene to transgenic plants and to assess the effect of the introduced gene on resistance to microbial pathogens by comparison with control plants.

The first and most important step in the strategy for disease control described above is to identify, characterise and describe a protein with strong anti-microbial activity. In recent years, many different plant proteins with anti-microbial and/or antifungal activity have been identified and described. These proteins have been categorised into several classes according to either their presumed mode of action and/or their amino acid sequence homologies. These classes include the following: chitinases (Roberts, W.K. *et al.* [1986] *Biochim. Biophys. Acta* 880:161-170); β-1,3-glucanases (Manners, J.D. *et al.* [1973] *Phytochemistry* 12:547-553); thionins (Bolmann, H. *et al.* [1988] *EMBO J.* 7:1559-1565 and Fernadez de Caleya, R. *et al.* [1972] *Appl. Microbiol.* 23:998-1000); permatins (Roberts, W. K. *et al.* [1990] *J. Gen. Microbiol.* 136:1771-1778 and Vigers, A.J. *et al.* [1991] *Mol. Plant-Microbe*

Interact. 4:315-323); ribosome-inactivating proteins (Roberts, W. K. *et al.* [1986] *Biochim. Biophys. Acta* 880:161-170 and Leah, R. *et al.* [1991] *J. Biol. Chem.* 266:1564-1573); plant defensins (Terras, F. R. G. *et al.* [1995] *The Plant Cell* 7:573-588); chitin binding proteins (De Bolle, M.F.C. *et al.* [1992] *Plant Mol. Biol.* 22:1187-1190 and Van Parijs, J. *et al.* [1991] *Planta* 183:258-264); thaumatin-like, or osmotin-like proteins (Woloshuk, C.P. *et al.* [1991] *The Plant Cell* 3:619-628 and Hejgaard, J. [1991] *FEBS Letts.* 291:127-131); PR1-type proteins (Niderman, T. *et al.* [1995] *Plant Physiol.* 108:17-27.) and the non-specific lipid transfer proteins (Terras, F.R.G. *et al.* [1992] *Plant Physiol.* 100:1055-1058 and Molina, A. *et al.* [1993] *FEBS Letts.* 3166:119-122). Another class of anti-microbial proteins from plants is the knottin or knottin-like anti-microbial proteins (Cammue, B.P.A. *et al.* [1992] *J. Biol. Chem.* 67:2228-2233; Broekaert W.F. et.al. (1996) *Crit. Rev. in Plant Sci.* (In press)). In addition, plants are not the sole source of anti-microbial proteins and there are many reports of the isolation of anti-microbial proteins from animal and microbial cells (reviewed in Gabay, J.E. [1994] *Science* 264:373-374 and in "Anti-microbial peptides" [1994] *CIBA Foundation Symposium 186*, John Wiley and Sons Publ., Chichester, UK).

There is some evidence that the ectopic expression of genes encoding proteins that have *in vitro* anti-microbial activity in transgenic plants can result in increased resistance to microbial pathogens. Examples of this engineered resistance include transgenic plants expressing genes encoding: a plant chitinase, either alone (Broglie, K. *et al.* [1991] *Science* 254:1194-1197) or in combination with a β-1,3-glucanase (Van den Elzen, P.J.M. *et al.* [1993] *Phil. Trans. Roy. Soc.* 342:271-278); a plant defensin (Terras, F.R.G. *et al.* [1995] *The Plant Cell* 7:573-588); an osmotin-like protein (Liu, D. *et al.* [1994] *Proc. Natl. Acad. Sci. USA* 91:1888-1892); a PR1-class protein (Alexander, D. *et al.* [1993] *Proc. Natl. Acad. Sci. USA* 90:7327-7331) and a ribosome-inactivating protein (Logemann, J. *et al.* [1992] *Bio/Technology* 10:305-308).

Although the potential use of anti-microbial proteins for engineering disease resistance in transgenic plants has been described extensively, there are other applications which are worthy of mention. Firstly, highly potent anti-microbial proteins can be used for the control of plant disease by direct application (De Bolle,

M.F.C. *et al.* [1993] in *Mechanisms of Plant Defence Responses*, B. Fritig and M. Legrand eds., Kluwer Acad. Publ., Dordrecht, NL, pp. 433-436). In addition, anti-microbial peptides have potential therapeutic applications in human and veterinary medicine. Although this has not been described for peptides of plant origin it is being actively explored with peptides from animals and has reached clinical trials (Jacob, L. and Zasloff, M. [1994] in "Anti-microbial Peptides", *CIBA Foundation Symposium 186*, John Wiley and Sons Publ., Chichester, UK, pp. 197-223).

The invention described herein constitutes a previously undiscovered and thus novel protein with anti-microbial activity. This protein can be isolated from *Macadamia integrifolia* (Mi) plants. *Macadamia integrifolia* belongs to the family Proteaceae. *M. integrifolia*, also known as Bauple Nut or Queensland Nut, is considered by some to be the world's best edible nut. For this reason it is extensively cultivated commercially, both in Australia and overseas (Williams, Keith A. W., *Native Plants (Queensland)*, Volume II, 1984, published by Keith A. W. Williams and printed by Printcraft of Newstead, Qld, Australia).

SUMMARY OF THE INVENTION

According to a first embodiment of the invention, there is provided an anti-microbial protein selected from:

(i) a protein having an amino acid sequence according to the sequence shown in Figure 4;

(ii) a homologue of (i);

(iii) a protein having at least 40% identity with (i) and which has essentially the same anti-microbial activity as (i);

(iv) a polypeptide fragment of (i), (ii) or (iii) which has essentially the same anti-microbial activity as (i);

(v) a protein having a 3-dimensional shape essentially the same as (i) as depicted in Figure 8;

(vi) a protein with a relative cysteine spacing of nX-C-3X-C-12X-C-3X-C-nX wherein X is any amino acid residue, n is any integer from 10-16, and C is cysteine; and

(vii) a protein with 40% or more positively charged residues having positions relative to cysteine residues of #-X-#-9X-C-X-#-#-C-X-#-#-

9X-C-X-X-#-C-X-#-#-X-X-#-X-X-#-#-X-X-#-# wherein X is any amino acid residue, C is cysteine and # is an amino acid residue with a positively charged side chain.

According to a second embodiment of the invention, there is provided an isolated DNA encoding a protein according to the first embodiment

According to a third embodiment of the invention, there is provided a DNA construct which includes a DNA according to the second embodiment operatively linked to elements for the expression of said encoded protein.

According to a fourth embodiment of the invention, there is provided a transgenic plant harbouring a DNA construct according to the third embodiment.

According to a fifth embodiment of the invention, there is provided reproductive material of a transgenic plant according to the fourth embodiment.

According to a sixth embodiment of the invention, there is provided a composition comprising an anti-microbial protein according to the first embodiment together with an agriculturally acceptable carrier diluent or excipient.

According to a seventh embodiment of the invention, there is provided a composition comprising an anti-microbial protein according to the first embodiment together with an pharmaceutically acceptable carrier diluent or excipient.

According to an eighth embodiment of the invention, there is provided a method of controlling microbial infestation of a plant, the method comprising:

i) treating said plant with an anti-microbial protein according to the first embodiment or a composition according to the sixth embodiment; or

ii) introducing a DNA construct according to the third embodiment into said plant.

According to a ninth embodiment of the invention, there is provided a method of controlling microbial infestation of a mammalian animal, the method comprising treating the animal with an anti-microbial protein according to the first embodiment or a composition according to the seventh embodiment.

According to a tenth embodiment of the invention, there is provided a method of preparing an anti-microbial protein, which method comprises the steps of:

a) designing an amino acid sequence having substantially the same distribution of positively charged residues relative to cysteine residues

as the amino acid sequence depicted in Figure 4;

b) synthesising a protein comprising said amino acid sequence; and

c) if necessary, forming disulphide linkages between said cysteine residues.

Other embodiments of the invention include methods for producing anti-microbial protein.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1a shows the results of cation-exchange chromatography of the basic protein fraction with the results of a bioassay for anti-microbial activity shown for fractions in the region of MiAMP2 elution. Figure 1b shows the results of including 1 mM Ca^{2+} in a parallel bioassay of fractions from the cation-exchange separation.

Figure 2 shows a reversed-phase HPLC profile of highly inhibitory fractions containing MiAMP2 from the cation-exchange separation in Figure 1a/b together with % inhibition exhibited by HPLC fractions.

Figure 3 shows the results of a mass spectrometric analysis of MiAMP2.

Figure 4 shows the amino acid sequence of MiAMP2.

Figure 5 shows a synthetic nucleotide sequence which can be used for the expression and secretion of MiAMP2 *in planta*.

Figure 6 shows the alignment of MiAMP2 with other proteins showing significant homology.

Figure 7 portrays a series of secondary structure predictions for MiAMP2.

Figure 8 shows a crude three-dimensional representation of MiAMP2 to illustrate how the spacing of positively charged residues in helical regions of the molecule will affect positioning of the side chains on one face of the helix.

BEST MODE AND OTHER MODES FOR CARRYING OUT THE INVENTION

The following abbreviations are used hereafter:

EDTA	ethylenediaminetetraacetic acid
MeCN	methyl cyanide (acetonitrile)
Mi	*Macadamia integrifolia*
MiAMP2	*Macadamia integrifolia* anti-microbial protein 2
ND	not determined
PCR	polymerase chain reaction

PMSF	phenylmethylsulphonyl fluoride
SDS-PAGE	sodium dodecylsulphate polyacrylamide gel electrophoresis
TFA	trifluoroacetate

The present inventors have identified a new class of anti-microbial proteins. A prototype protein can be isolated from seeds of *Macadamia integrifolia* (hereafter Mi). The invention thus provides anti-microbial protein *per se* as well as DNA sequences encoding anti-microbial protein.

The invention also provides an amino acid sequence of the prototype anti-microbial protein. From this sequence, a nucleotide sequence of DNA encoding the protein can be derived by reverse translating the amino-acid sequence. DNA having a nucleotide sequence encoding anti-microbial protein can then be synthesised chemically (and/or enzymically) or isolated from plant tissue of the plants in question (i.e. Mi) using standard cloning methods as described in laboratory manuals such as *Current Protocols in Molecular Biology* (copyright 1987-1995 edited by Ausubel F. M. *et al.* and published by John Wiley & Sons, Inc. printed in the USA).

The anti-microbial protein *per se* will manifest a particular three-dimensional structure which may be determined using X-ray crystallography or nuclear magnetic resonance techniques. This structure will be responsible in large part for the anti-microbial activity of the protein. From the sequence of the protein, it is also possible to make predictions concerning possible conformations and structural motifs (secondary structure) that will likely be exhibited by the protein. It will be appreciated that one skilled in the art could take a protein with known structure and alter the sequence significantly and yet retain the overall three-dimensional shape and anti-microbial activity of the protein. One aspect of the structure that most likely could not be altered without seriously affecting the structure (and, therefore, the activity of the protein) is the content and spacing of the cysteine residues since this would disrupt the formation of disulfide bonds which are critical to maintaining the overall structure of the protein and making the protein more stable. In particular, it is essential that each pair of cysteine residues be separated by exactly 3 residues.

The distribution of positive (and negative) charges on the various surfaces of the protein will also have critical roles in determining the structure and activity of the protein. In particular, the distribution of charged residues in an α-helical region of

a protein can result in charges lying on one face of the helix. The spacing which is required for residues to be positioned on one face of the helix can easily be determined by one skilled in the art using a helical wheel plot with the sequence of interest. A helical wheel plot uses the fact that in α-helices each turn of the helix is composed of 3.6 residues on the average. This number translates to 100 degrees of rotational translation per residue making it possible to construct a plot showing the distribution of side chains in a helical region. Figure 8 was produced from a helical wheel plot and shows how the spacing of charged residues can lead to most of the positively charged side chains being localised on one face of the helix.

It will be appreciated by one of skill in the art that positive charges are typically conferred on the face of a protein by arginine, lysine and histidine residues.

DNA sequences coding for these proteins can be deduced using standard codon tables. Using this finite number of possible DNA sequences, one skilled in the art could design oligonucleotide probes which can be used to isolate the actual gene(s) as well as control sequences from the plants. It is also possible to chemically synthesise the gene using a standard DNA synthesis instrument (such as a Beckman Oligo1000 instrument) together with known techniques for constituting synthesised gene fragments into a whole gene (see *Current Protocols in Molecular Biology, supra*).

This gene, under control of a constitutive or inducible promoter, can then be cloned into a biological system which allows expression of the protein. Transformation methods allowing for the protein to be expressed in a variety of systems are known. The protein can thus be expressed in any suitable system for the purpose of producing the protein for further use. Suitable hosts for the expression of this protein include *E. coli*, fungal cells, insect cells, mammalian cells, and plants. Standard methods for expressing proteins in such hosts are described in a variety of texts including section 16 (Protein Expression) of *Current Protocols in Molecular Biology, (supra)*.

Plant cells may be transformed with DNA constructs of the invention according to a variety of known methods (*Agrobacterium*, Ti plasmids, electroporation, micro-injections, micro-projectile gun, and the like). For expression in plants, the DNA sequence encoding the *Macadamia integrifolia* anti-microbial protein number 2

(hereafter MiAMP2) can be used in conjunction with a DNA sequence encoding a preprotein from which the mature protein is produced. This preprotein may contain a native or synthetic signal peptide sequence which would target the protein to a particular cell compartment (e.g. the apoplast or the vacuole). These coding sequences can be ligated to a plant promoter sequence that will ensure strong expression in plant cells. This promoter sequence might ensure strong constitutive expression of the protein in most or all plant cells, it may be a promoter which ensures expression in specific tissues or cells that are susceptible to microbial infection and it may also be a promoter which ensures strong induction of expression during the infection process. These types of gene cassettes will also include a transcription termination and polyadenylation sequence 3' of the MiAMP2 coding region to ensure efficient production and stabilisation of the mRNA encoding the MiAMP2. It is possible that efficient expression of MiAMP2 might be facilitated by inclusion of its DNA sequence into a sequence encoding a much larger protein which is processed *in planta* to produce one or more active MiAMP2 molecules. Gene cassettes encoding MiAMP2 as described above can then be expressed in plant cells using two common methods. Firstly, the gene cassettes can be ligated into binary vectors carrying, i) left and right border sequences that flank the T-DNA of the *Agrobacterium tumefaciens* Ti plasmid, ii) a suitable selectable marker gene for the selection of antibiotic resistant plant cells, iii) origins of replication that function in either *A. tumefaciens* or *Escherichia coli* and iv) antibiotic resistance genes that allow selection of plasmid-carrying cells of *A. tumefaciens* and *E. coli*. This binary vector carrying the chimaeric MiAMP2 encoding gene can be introduced by either electroporation or triparental mating into *A. tumefaciens* strains carrying disarmed Ti plasmids such as strains LBA4404, GV3101, and AGL1 or into *A. rhizogenes* strains such as R4 or NCCP1885. These *Agrobacterium* strains can then be co-cultivated with suitable plant explants or intact plant tissue and the transformed plant cells and/or regenerants selected using antibiotic resistance.

The expression of the MiAMP2 protein in the transgenic plants can be detected using either antibodies raised to the protein or using anti-microbial bioassays. A second method of gene transfer to plants can be achieved by direct insertion of the gene in target plant cells. For example, the MiAMP2 encoding gene cassette can be

co-precipitated onto gold or tungsten particles along with a plasmid encoding a chimaeric gene for antibiotic resistance in plants. The tungsten particles can be accelerated using a fast flow of helium gas and the particles allowed to bombard a suitable plant tissue. This can be an embryogenic cell culture, a plant explant, a callus tissue or cell suspension or an intact meristem. Plants can be recovered using the antibiotic resistance gene for selection and antibodies used to detect plant cells expressing the MiAMP2 protein. These and other related methods for the expression of the MiAMP2 in plants are described in *Plant Molecular Biology* (2nd ed., edited by Gelvin, S.B. and Schilperoort, R.A.,© 1994, published by Kluwer Academic Publishers, Dordrecht, The Netherlands)

Both monocotyledonous and dicotyledonous plants can be transformed and regenerated. Examples of genetically modified plants include maize, banana, peanut, field peas, sunflower, tomato, canola, tobacco, wheat, barley, oats, potato, soybeans, cotton, carnations, roses, sorghum. These, as well as other agricultural plants can be transformed with the anti-microbial genes such that they would exhibit a greater degree of resistance to pathogen attack. Alternatively the proteins can be used for the control of diseases by topological application.

The invention also relates to application of anti-microbial protein in the control of pathogens of mammals including humans. The protein can be used either in topological or intravenous applications for the control of microbial infections.

As indicated above in the description of the tenth embodiment, the invention includes within its scope the preparation of anti-microbial proteins based on the prototype Mi protein. From the Mi protein amino acid sequence, a new sequence can be designed which retains a distribution of positively charged residues relative to cysteine residues substantially the same as the Mi protein. The new sequence can be synthesised or expressed from a gene encoding the sequence in an appropriate host cell. Suitable methods for such procedures have been described above. Expression of the new protein in a genetically engineered cell will typically result in a product having a correct three-dimensional structure, including correctly formed disulphide linkages between cysteine residues. However, even if the protein is chemically synthesised, methods are well known in the art for further processing of the protein to give the desired three-dimensional structure.

Macadamia integrifolia anti-microbial protein number 2

As indicated above, a new potent anti-microbial protein has been identified in the seeds of Mi. The protein factor is called MiAMP2. The protein is highly basic with a predicted pI value of 11.55 and contains 4 cysteine residues which are presumed to be important in stabilising the three-dimensional structure of the protein through the formation of disulfide bonds. Additionally, the relative molecular mass of the protein has been determined by mass spectrometry which shows it to be 6216.8 +/- 1 Da. The amino acid sequence shares some homology with previously described proteins in sequence databases (Swiss Prot and Non-redundant databases) searched using the BLASTP algorithm (*Altschul, S.F. et al.* [1990] *J. Mol. Biol.* 215:403). In particular, 25 of 47 residues in MiAMP2 (53%) align with a portion of cocoa vicilin and 18 of 47 (38%) align with cotton vicilin in two non-adjacent regions. These vicilins are much larger proteins than MiAMP2 with 452 and 499 amino acids, respectively (see Figure 6). MiAMP2 also shares some homology with MBP-1 anti-microbial protein from maize (11 of 47, 23%, of MiAMP2 residues align with the 33 residue MBP-1 protein) (Duvick, J.P. et al. (1992) *J Biol Chem* 267:18814-20). The alignments in Figure 6 show the similarity in cysteine spacing between MiAMP2 and the vicilin molecules. In Figure 6, cysteine residues are highlighted with inverted text and positively charged residues in other proteins which align with positively charged residues of MiAMP2 are shown in shaded boxes. Residues which are identical to MiAMP2 are in bold type.

The MiAMP2 protein shows a wide range of anti-fungal activity. MiAMP2 shows very significant inhibition of fungal growth at concentrations as low as 2 μg/ml for some of the pathogens/microbes against which the protein was tested. Thus it can be used to provide protection against several plant diseases. MiAMP2 can be used as a fungicide or antibiotic by application to plant parts. The protein can also be used to inhibit growth of pathogens by expressing it in whole transgenic plants. The protein can be used for the control of human pathogens by topological application or intravenous injection. One characteristic of the protein is that inhibition of some microbes is suppressed by the presence of Ca^{2+} (1 mM) (see Table I)

The MiAMP1 protein also has potential as an insect control agent. Since the protein is extremely basic (pI > 11.5), it would maintain a positive charge even in

the highly alkaline environment of an insect gut enabling it to interact with negatively charged structures within the gut. This interaction may lead to inefficient feeding, slowing of growth, and possibly death of the insect.

Non-limiting examples of the invention follow.

EXAMPLE 1

Extraction of Basic Protein from *Macadamia integrifolia* Seeds

Twenty five kilograms of Mi nuts (purchased from the Macadamia Nut Factory, Queensland, Australia) were ground in a food processor (The Big Oscar, Sunbeam) and the resulting meal was extracted for 2-4 hours at 4°C with 50 L of an ice-cold extraction buffer containing 10 mM NaH_2PO_4, 15 mM Na_2HPO_4, 100 mM KCl, 2 mM EDTA, 0.75% polyvinylpolypyrrolidone, and 0.5 mM phenylmethylsulfonyl fluoride (PMSF). The resulting homogenate was run through a kitchen strainer to remove larger particulate material and then further clarified by centrifugation (4000 rpm for 15 min) in a large capacity centrifuge. Solid ammonium sulphate was added to the supernatant to obtain 30% relative saturation and the precipitate allowed to form overnight with stirring at 4°C. Following centrifugation at 4000 rpm for 30 min, the supernatant was taken and ammonium sulphate added to achieve 70% relative saturation. The solution was allowed to precipitate overnight and then centrifuged at 4000 rpm for 30 min in order to collect the precipitated protein fraction. The precipitated protein was resuspended in a minimal volume of extraction buffer and centrifuged once again (13,000 rpm x 30 min) to remove the undissolved portion. After dialysis (10 mM ethanolamine pH 9.0, 2 mM EDTA and 1 mM PMSF) to remove residual ammonium sulphate, the protein solution was passed through a Q-Sepharose Fast Flow column (5 x 12 cm) previously equilibrated with 10 mM ethanolamine (pH 9), 2 mM in EDTA).

The collected flowthrough from this column represents the basic (pI $\geq$9) protein fraction of the seeds. This fraction was further purified as described in Example 3.

EXAMPLE 2

Antifungal and Antibacterial Activity Assays

In general, bioassays to assess antifungal and antibacterial activity were carried

out in 96-well microtitre plates. Typically, the test organism was suspended in a synthetic growth medium consisting of K_2HPO_4 (2.5 mM), $MgSO_4$ (50 μM), $CaCl_2$ (50 μM), $FeSO_4$ (5 μM), $CoCl_2$ (0.1 μM), $CuSO_4$ (0.1 μM), Na_2MoO_4 (2 μM), H_3BO_3 (0.5 μM), KI (0.1 μM), $ZnSO_4$ (0.5 μM), $MnSO_4$ (0.1 μM), glucose (10 g/L), asparagine (1 g/L), methionine (20 mg/L), myo-inositol (2 mg/L), biotin (0.2 mg/L), thiamine-HCl (1 mg/L) and pyridoxine-HCL (0.2 mg/L). The test organism consisted of bacterial cells, fungal spores (50,000 spores/ml) or fungal mycelial fragments (produced by blending a hyphal mass from a culture of the fungus to be tested and then filtering through a fine mesh to remove larger hyphal masses). Fifty microlitres of the test organism suspended in medium was placed into each well of the microtitre plate. A further 50 μl of the test anti-microbial solution was added to appropriate wells. To deal with well-to-well variability in the bioassay, 4 replicates of each test solution were done. Sixteen wells from each 96-well plate were used as controls for comparison with the test solutions.

Unless otherwise stated, the test organism used was *Phytophthora cryptogea* and incubation was at 25°C for 48 hours. All fungi including yeast were grown at 25°C. *E. coli* were grown at 37°C and other bacteria were bioassayed at 28°C. Percent growth inhibition was measured by following the absorbance at 600 nm of growing cultures over various time intervals and is defined as 100 times the ratio of the average change in absorbance in the control wells minus the change in absorbance in the test well divided by the average change in absorbance at 600 nm for the control wells (i.e. [(avg change in control wells - change in test well) / (avg change in control wells)] x 100). Typically, measurements were taken at 24 hour intervals and the period from 24-48 hours was used for %Inhibition measurements.

Table 1 shows the IC_{50} value of pure MiAMP2 against various fungi and bacteria. In the table, the ">50" indicates that concentrations higher than 50 μg/ml were not tested. The abbreviation "ND" indicates that the test was not performed or that results could not be interpreted. The anti-microbial protein was also tested in the presence of 1 mM Ca^{2+} in the test medium and the IC_{50} values for these tests are given in the right-hand column. As can be seen in the table, the inhibitory activity of MiAMP2 is greatly reduced (although not eliminated) in the presence of Ca^{2+}.

Table 1

Concentrations of MiAMP2 at which 50% inhibition of growth was observed

Organism	IC_{50} (μg/ml)	IC_{50} + Ca^{2+} (μg/ml)
Alternaria helianthi	5-10	ND
Candida albicans	>50	>50
Ceratocystis paradoxa	20-50	>50
Cercospora nicotianae	5-10	5-10
Clavibacter michiganensis	50	>50
Chalara elegans	2-5	10-20
Fusarium oxysporum	10	20-50
Sclerotinia sclerotiorum	20-50	>50
Phytophthora cryptogea	5-10	10-25
Phytophthora parasitica nicotiana	10-20	>50
Verticillium dahliae	5-10	>50
Ralstonia solanacearum	>50	>50
Pseudomonas syringae tabaci	>50	>50
Saccharomyces cerevisiae	20-50	>50
Escherichia coli	>50	>50

EXAMPLE 3

Purification of Anti-microbial Protein from *Macadamia integrifolia* Basic Protein Fraction

The starting material for the isolation of the Mi anti-microbial protein was the basic fraction extracted from the mature seeds as described above in Example 1. This protein was further purified by cation exchange chromatography as shown in Figure 1a.

About 4 g of the basic protein fraction dissolved in 20 mM sodium succinate (pH 4) was applied to an S-Sepharose High Performance column (5 X 60 cm) (Pharmacia) previously equilibrated with the succinate buffer. The column was eluted at 17 ml/min with a linear gradient of 20 L from 0 to 2 M NaCl in 20 mM sodium succinate (pH 4). The eluate was monitored for protein by on-line measurement of the absorbance at 280 nm and collected in 200 ml fractions. Portions of each fraction were subsequently tested in the antifungal activity assay against *Phytopthora cyrptogea* at a concentration of 100 μg/ml in the presence and absence of 1 mM Ca^{2+}. Results of bioassays are included in Figures 1a and 1b where the elution gradient is shown as a solid line and the shaded bars represent %Inhibition. The Figure 1a assays were conducted without added Ca^{2+} while 1 mM Ca^{2+} was included in the Figure 1b assays. Fractionation yielded a number of unresolved peaks eluting between 0.05 and 2 M NaCl. A peak eluting at about 16 hours into the separation (fractions 90-92) showed significant anti-microbial activity.

Fractions showing significant anti-microbial activity were further purified by reversed-phase chromatography. Aliquots of fractions 90-92 were loaded onto a Pep-S (C_2/C_{18}), column (25 x 0.93 cm) (Pharmacia) equilibrated with 95% H_2O/5% MeCN/0.1% TFA (=100%A). The column was eluted at 3 ml/min with a 240 ml linear gradient (80 min) from 100%A to 100%B (=5% H_2O/95% MeCN/0.1% TFA). Individual peaks were collected, vacuum dried three times in order to remove traces of TFA, and subsequently resuspended in 500 μl of milli-Q water (Millipore Corporation water purification system) for use in bioassays as described in Example 2. Figure 2 shows the HPLC profile of purified fraction 92 from the cation-exchange separation shown in Figure 1a/b. Protein elution was monitored at 214 nm. The acetonitrile gradient is shown by the straight line. Individual peaks were bioassayed for anti-microbial activity: the bars in Figure 2 show the inhibition corresponding to 15 μg/ml of material from each of the fractions. The active protein elutes at approximately 27 min (~30% MeCN/0.1%TFA) and is called MiAMP2.

EXAMPLE 4

Purity of Isolated MiAMP2

The purity of the isolated anti-microbial protein was verified by native SDS-PAGE followed by staining with coomassie blue protein staining solution.

Electrophoresis was performed on a 10-20% tricine gradient gel (Novex) as per the manufacturers recommendations (100 V, 1-2 hour separation time). Under these conditions the purified MiAMP2 migrates as a single discrete band (< 10 kDa in size). The detection of a single major band in the SDS-PAGE analysis together with single peaks eluting in the cation-exchange and reversed-phase separations (not shown), gives strong indication that the activity of the MiAMP2 was due to the purified protein alone and not to a minor contaminating component.

EXAMPLE 5

Mass Spectroscopic Analysis of MiAMP2

Purified MiAMP2 was submitted for mass spectroscopic analysis. Approximately 1 μg of protein in solution was used for testing. Analysis showed the protein to have a molecular weight of 6216.8 Da $\pm$ 1 Da (see Figure 3 in which mass/charge data, A, and a reconstruction of the molecular weight, B, are presented). Additionally, the protein was subjected to reduction of disulfide bonds with dithiothreitol and alkylation with 4-vinylpyridine. The product of this reduction/alkylation was then submitted for mass spectroscopic analysis and was shown to have gained 427 mass units (i.e. molecular weight was increased by approximately 4 X 106 Da). The gain in mass indicated that four 4-vinylpyridine groups had reacted with the reduced protein, demonstrating that the protein contains a total of 4 cysteine residues. The cysteine content has also been subsequently confirmed through amino acid sequencing (see Example 6).

EXAMPLE 6

Amino Acid Sequence of MiAMP2 Protein

Approximately 1 μg of the pure protein which had been reduced and alkylated was subjected to Automated Edman degradation N-terminal sequencing. In the first sequencing run, the sequence of the first 39 residues was determined. Subsequently, approximately 1 mg of MiAMP2 was reacted with Cyanogen Bromide which cleaved the protein on the C-terminal side of Methionine-26. The C-terminal fragment generated by the cleavage reaction was purified by reversed-phase HPLC and sequenced, yielding the remaining sequence of MiAMP2 (i.e. residues 27-47). The full amino acid sequence is presented in Figure 4. In Figure 4, cysteine residues are in bold type and underlined. Positively charged residues are underlined. Depending

on the number of disulfide bonds that are formed, the protein mass will range from 6215.6 to 6219.6 Da. This is in close agreement with the mass of 6216.8 ± 1 Da obtained by mass spectrometric analysis (Example 5). The measured mass closely approximates the predicted mass of MiAMP2 in a two-disulfide form as it is expected to be found but also approaches the predicted mass of a one-disulfide form of MiAMP2.

EXAMPLE 7

Synthetic DNA Sequence Coding for MiAMP2 with a leader peptide

Using standard codon tables it is possible to reverse-translate the protein sequence to obtain DNA sequences that will code for the antimicrobial protein. The software program MacVector 4.5.3 was used to enter the protein sequence and obtain a degenerate nucleotide sequence. A codon usage table for tobacco was referenced in order to pick codons that would be adequately represented in tobacco for purposes of obtaining high expression in this test plant. A 30 amino-acid leader peptide was also designed to ensure efficient processing of the signal peptide and secretion of the peptide extracellularly. For this purpose, the method of Von Hiejne was used to evaluate a series of possible leader sequences for probability of cleavage at the correct position [Von Hiejne, G.(1986) *Nucleic Acids Research* 14(11): 4683-4690]. In particular, the amino acid sequence M-A-W-F-H-V-S-V-C-N-A-V-F-V-V-I-I-I-I-M-L-L-M-F-V-P-V-V-R-G was found to give an optimal probability of correct processing of the signal peptide immediately following the G of the leader sequence. A 5' untranslated region from tobacco mosaic virus was also added to this synthetic gene to promote higher translational efficiency [Dowson, M.J., et al. (1994) *Plant Mol. Biol. Rep.* 12(4):347-357]. The synthetic gene also contains restriction sites at the 5' and 3' ends and immediately 5' of the start ATG for efficient cloning and subcloning procedures. Figure 5 shows a synthetic DNA sequence suitable for use in plant expression experiments. In this figure, the arrow shows where translation is initiated and the triangular symbol indicates the point of cleavage of the signal peptide.

EXAMPLE 8

Partial DNA Sequence encoding MiAMP2 Protein

With the reverse-translated nucleotide sequences, degenerate primers were made for use in PCR reactions with genomic DNA from Macadamia. Primer JPM17 sequence was 5' CAG CAG CAG TAT GAG CAG TG 3' and primer JPM20 sequence was 5'TTT TTC GTA (T/T)C(T/G) (G/T)C(T/G) TTC GCA 3'. Primers JPM17 and JPM20 were used in PCR reactions carried out for 30 cycles with 30 sec at 95°C, 1 min at 50°C, and 1 min at 72°C. PCR products with sizes close to those which were expected were directly sequenced (ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit from Perkin Elmer Corporation) after excising DNA bands from agarose gels and purifying them using a Qiagen DNA clean-up kit. Using this approach, we were able to amplify a ~80 bp fragment of DNA. Direct sequencing of this fragment yielded the nucleotide sequence corresponding to a portion of the amino acid sequence of the anti-microbial protein MiAMP2 (amino acids 7-39 Figure 4). The nucleotide sequence of the above-mentioned fragment excluding the primer sequences is 5' TCA GAA GCG CTG CCA ACG GCG CGA GAC AGA GCC ACG ACA CAT GCA AAT TTG TCA ACA ACG C3'. This sequence can be used for a variety of purposes including (but not limited to) screening of cDNA and genomic libraries for clones of MiAMP2 or design of specific primers for PCR amplification reactions.

EXAMPLE 9

Crude structure prediction of MiAMP2 Protein

Using sequence analysis algorithms, putative secondary structure motifs can be assigned to the protein. Five different algorithms were used to predict whether α-helices, β-sheets, or turns can occur in the protein with the sequence given in Figure 4 [DPM method: Deleage, G., and Roux, B. (1987) *Prot. Eng.* 1:289-294; SOPMA method: Geourjon, C., and Deleage, G. (1994) *Prot. Eng.* 7:157-164; Gibrat method: Gibrat, J.F., Garnier, J., and Robson, B.(1987) *J.Mol.Biol.* 198:425-443; Levin method: Levin, J.M., Robson, B., And Garnier, J. (1986) *FEBS Lett.* 205:303-308; PhD method: Rost, B., And Sander, C. (1994) *Proteins* 19:55-72]. Figure 7 shows the predicted locations of α-helices, β-sheets and turns. The following symbols have been used in Figure 7: C, coil; H, alpha helix;

E, beta sheet; and S, turn. Underlined residues are represented as helices in Figure 8 (see following paragraph).

It is clear from the secondary structure predictions that the protein is highly α-helical. Examination of the secondary-structure predictions show a clear preponderance of two α-helical regions broken by a stretch of about 5 residues. This is highly suggestive of a helix-turn-helix motif. If each of the regions which are predicted to form helices are plotted on a helical wheel diagram, a startling pattern of positive charge distribution on one face of each helix becomes evident (see Figure 8). It can also be seen that cysteine residues with a C-3X-C spacing will be aligned on one face of the helix. Since the cysteines are involved in disulfide bond formation, the cysteine side chains must form covalent bonds with the cysteines located on the other helical segment. When the helical segments are arranged in such a way as to bring the cysteine side chains into proximity with the cysteine side chains from the other helix, the result is a three-dimensional structure which is likely to represent the actual 3-D structure of the protein. This structure exhibits a remarkable distribution of positively charge residues on one face of the protein comprised of two helices held together by two disulfide bonds (see Figure 8 in which positive charges are shown as circles).

DATED THIS 20TH DAY OF DECEMBER 1996

COOPERATIVE RESEARCH CENTRE

FOR TROPICAL PLANT PATHOLOGY

By their Patent Attorneys

CULLEN & CO

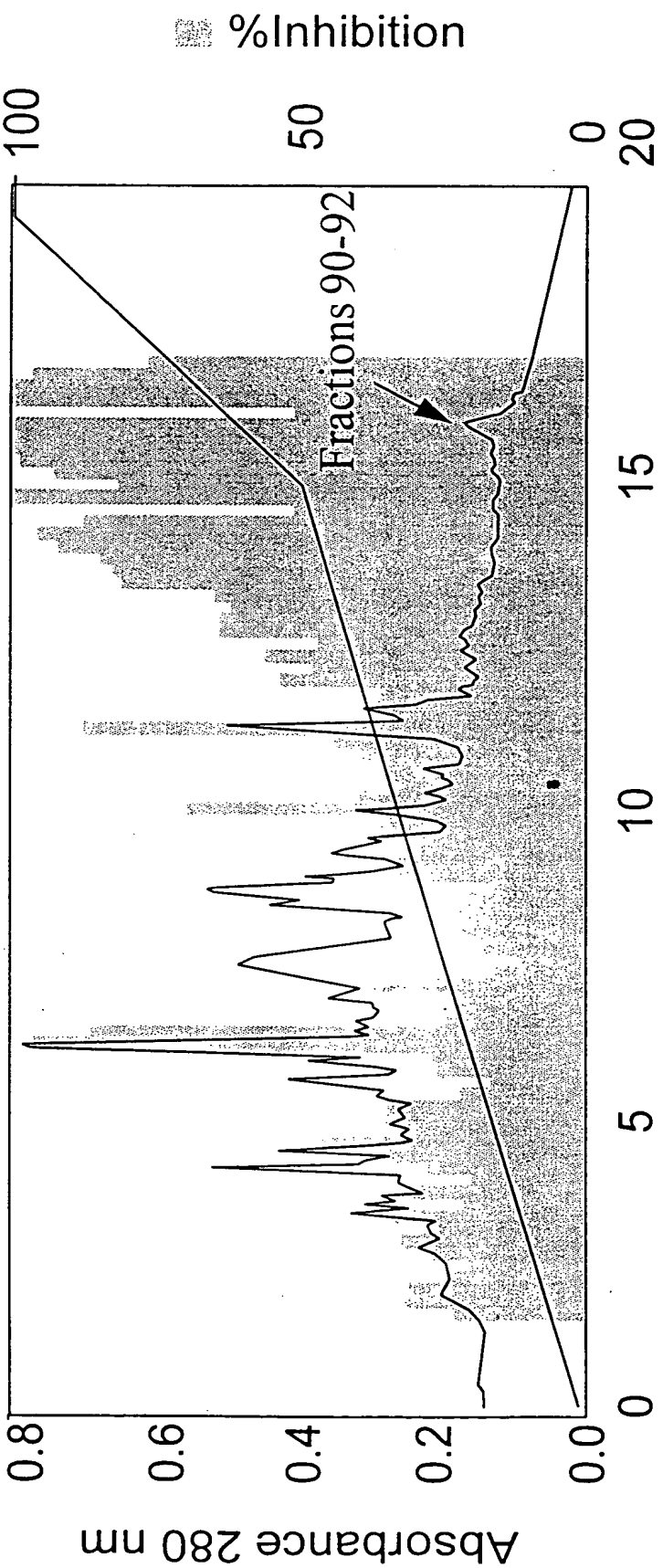
%Inhibition
100
50
0
20
Fractions 90-92
15
10
5
0
Hours
0.8
0.6
0.4
0.2
0.0
Absorbance 280 nm

Fig. 1a

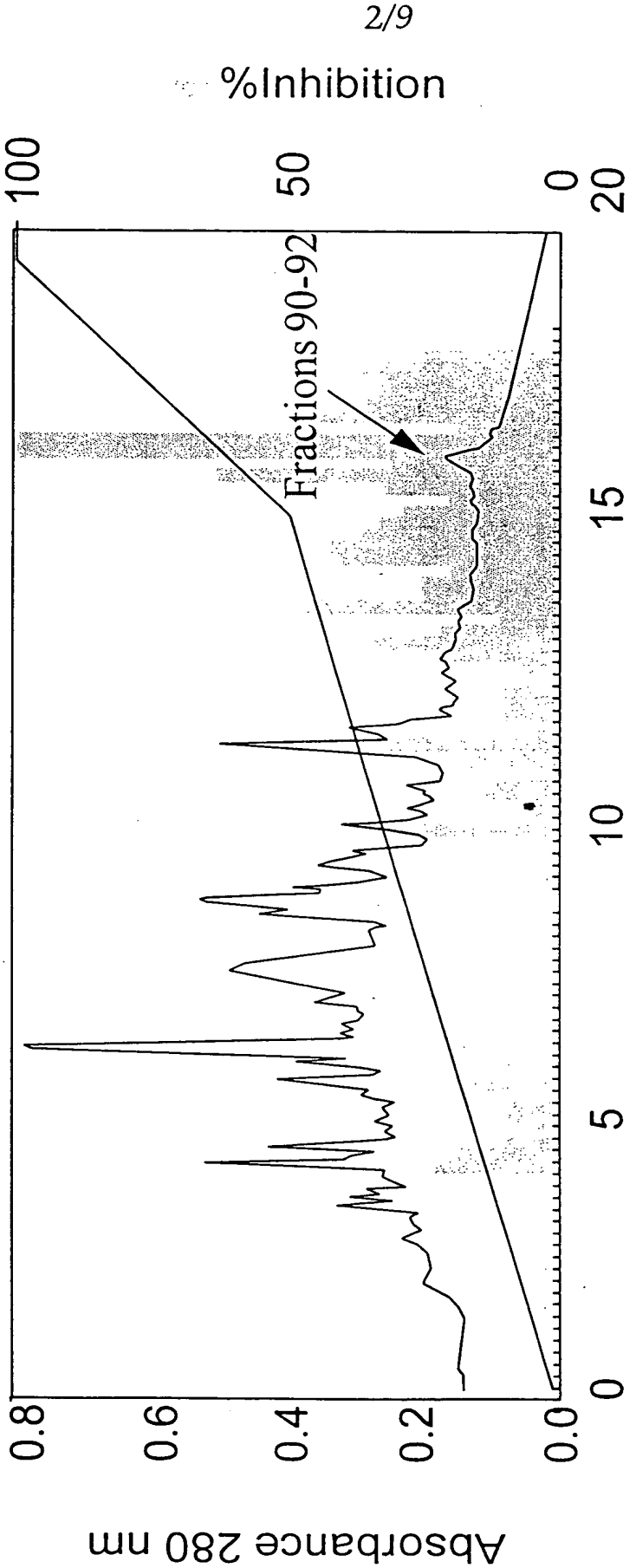
%Inhibition
100
50
0
20
15
10
5
0
Hours
Fractions 90-92
0.8
0.6
0.4
0.2
0.0
Absorbance 280 nm

Fig. 1b

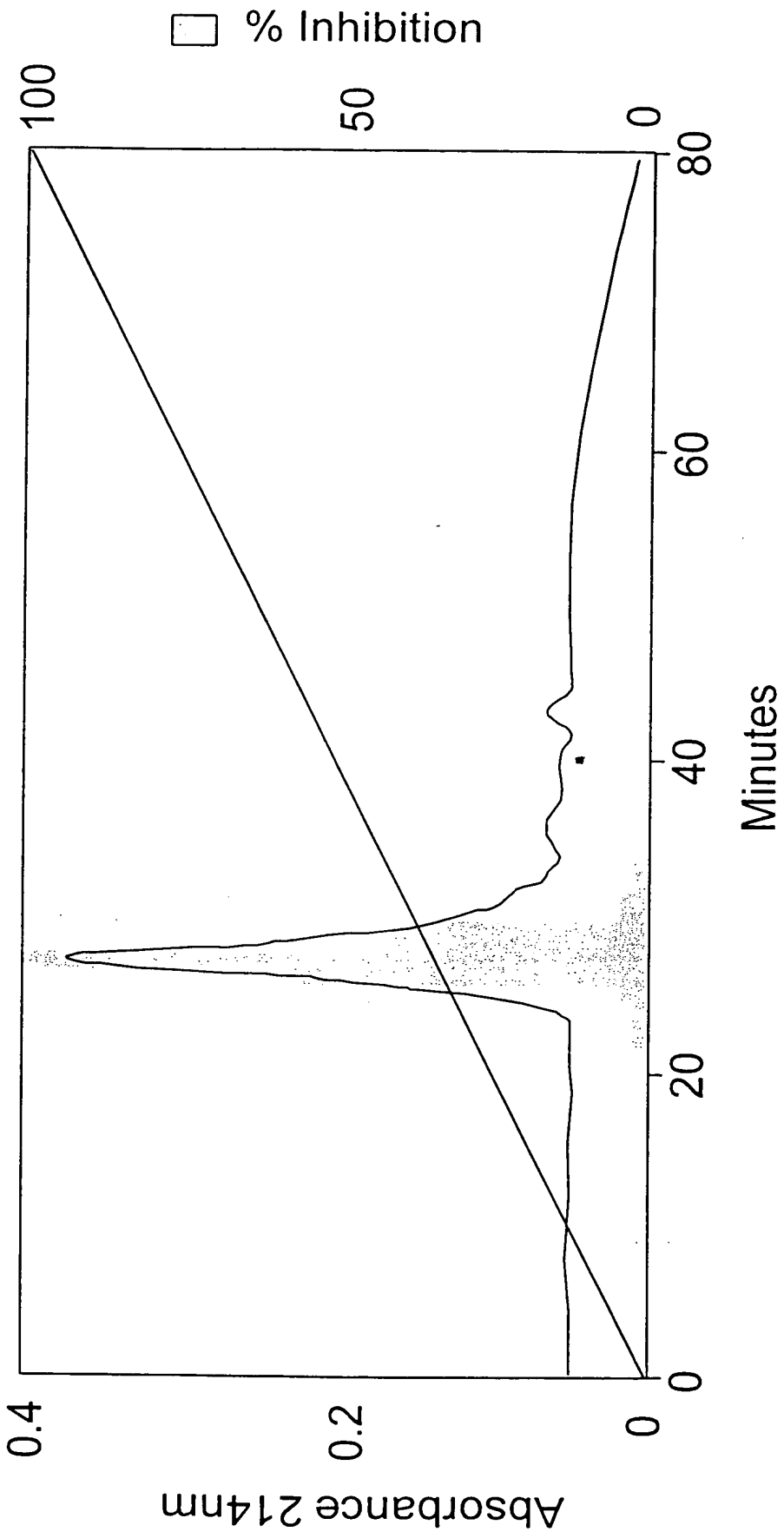
% Inhibition
100
50
0
80
60
40
20
0
Minutes
0.4
0.2
0
Absorbance 214nm

Fig. 2

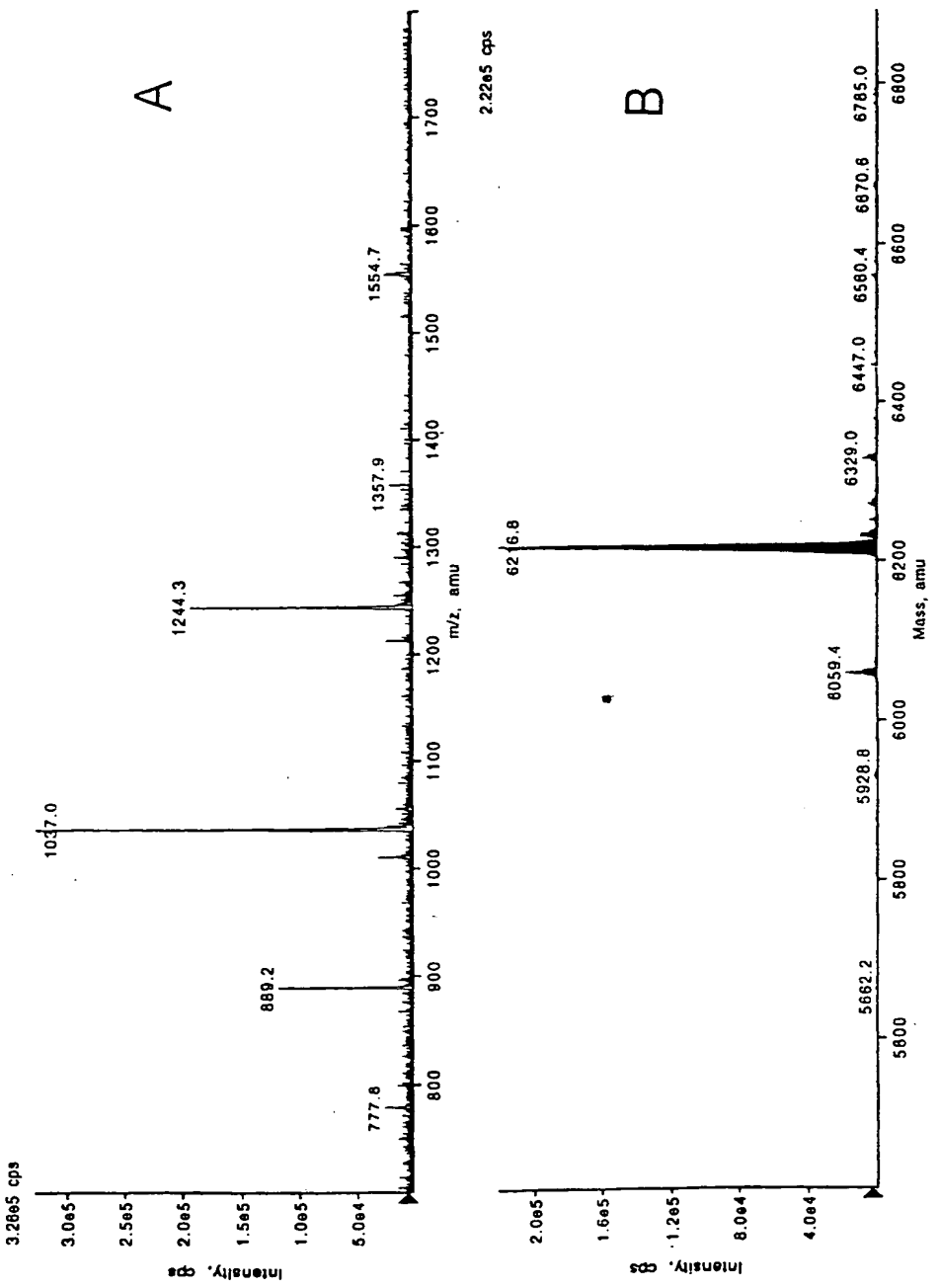
A
3.28e5 cps
3.0e5
2.5e5
2.0e5
1.5e5
1.0e5
5.0e4
Intensity, cps
777.8
889.2
1037.0
1244.3
1357.9
1554.7
800
900
1000
1100
1200
1300
1400
1500
1600
1700
m/z, amu
B
2.22e5 cps
2.0e5
1.6e5
1.2e5
8.0e4
4.0e4
Intensity, cps
5662.2
5928.8
6059.4
6216.8
6329.0
6447.0
6560.4
6670.8
6785.0
5800
5800
6000
6200
6400
6600
6800
Mass, amu

Fig. 3


1 10 20 30 40 47
RQRDPQQQYE QCQKRCQRRE TEPRHMQICQ QRCERRYEKE KRKQQKR

Fig. 4

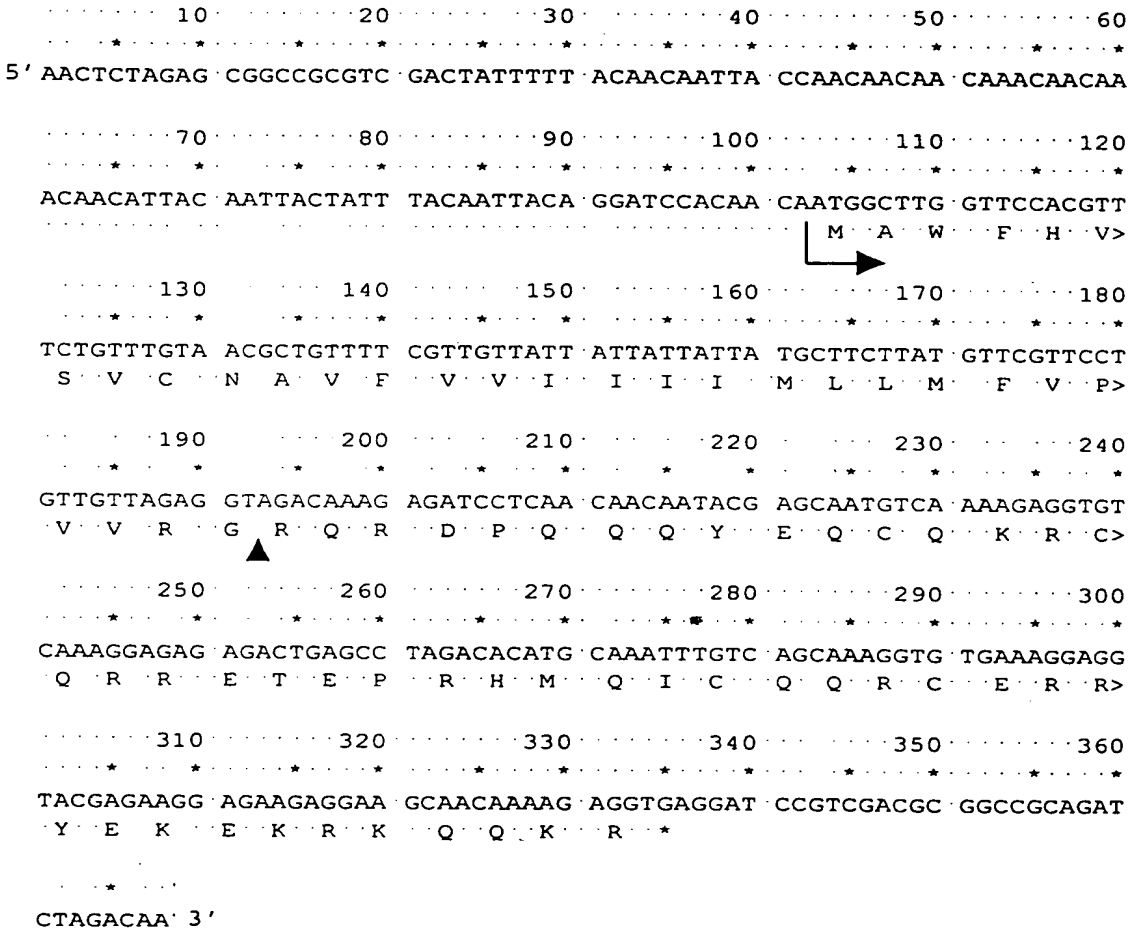
10 20 30 40 50 60
5' AACTCTAGAG CGGCCGCGTC GACTATTTTT ACAACAATTA CCAACAACAA CAAACAACAA
70 80 90 100 110 120
ACAACATTAC AATTACTATT TACAATTACA GGATCCACAA CAATGGCTTG GTTCCACGTT
M A W F H V>
130 140 150 160 170 180
TCTGTTTGTA ACGCTGTTTT CGTTGTTATT ATTATTATTA TGCTTCTTAT GTTCGTTCCT
S V C N A V F V V I I I I M L L M F V P>
190 200 210 220 230 240
GTTGTTAGAG GTAGACAAAG AGATCCTCAA CAACAATACG AGCAATGTCA AAAGAGGTGT
V V R G R Q R D P Q Q Q Y E Q C Q K R C>
250 260 270 280 290 300
CAAAGGAGAG AGACTGAGCC TAGACACATG CAAATTTGTC AGCAAAGGTG TGAAAGGAGG
Q R R E T E P R H M Q I C Q Q R C E R R>
310 320 330 340 350 360
TACGAGAAGG AGAAGAGGAA GCAACAAAAG AGGTGAGGAT CCGTCGACGC GGCCGCAGAT
Y E K E K R K Q Q K R *
CTAGACAA 3'

Fig. 5

```
Cotton vicilin:                                   MVRNKSACVVLLFSLFLSFGLLCSAKDFPGRR-  32
Cocoa vicilin:                                    MVISKSPFIVLIFSLLLSFALLCSGVSAYGRKQ  33

Cotton vicilin: GDDDPPKRYEDCRRRCEWDTRGQKEQQQCEESCKSQYGEKDQQQRHR                    79
Cocoa vicilin:  RERDPRQQYEQCQRRCESEATEEREQEQCEQRCEREYKEQQRQQEEE                    80
MIAMP2:         RQRDPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQKR*                   47

Cotton vicilin:   PEDPQRRYEECQQECRQQEERQRP--QCQQRCLKRFEQEQQQSQRQ                   123
Cocoa vicilin:    ---LQRQYQQCQGRCQEQQQGQREQQQCQRKCWEQYKEQERGEHEN                   123

Cotton vicilin:   -------FQECQQHCHQQEQRPERKQQCVRECRERYQENPWRRERE                   163
Cocoa vicilin:    YHNHKKNRSEEEEGQQRNNPYYFPKRRSFQTRFRDEEGNRKILQRFA                  170

Cotton vicilin:   EEAEEEETEEGEQEQSHNPFHFHRRSFQSRFREEHGNFRVLQRFASR                  210
Cocoa vicilin:    ENSPPLKGINDYRLAMFEANPNTFILPHHCDAEAIYFVTNGKGTITF                  217

Cotton vicilin:   HPILRGINEFRLSILEANPNTFVLPHHCDAEKIYLVTNGRGTLTFLT                  257
Cocoa vicilin:    VTHENKESYNVQRGTVVSVPAGSTVYVVSQDNQEKLTIAVLALPVNS                  264

Cotton vicilin:   HENKESYNVVPGVVVRVPAGSTVYLANQDNKEKLIIAVLHRPVNNPR                  304
Cocoa vicilin:    PGKYELFFPAGNNKPESYYGAFSYEVLETVFNTQREKLEEILEEQRG                  311

Cotton vicilin:   QFEEFFPAGSQRPQSYLRAFSREILEPAFNTRSEQLDELFGGRQSHR                  351
Cocoa vicilin:    QKRQQGQQGMFRKAKPEQIRAISQQATSPRHRGGERLAINLLSQSPV                  358

Cotton vicilin:   RQQGQGMFRKASQEQIRALSQEATSPREKSGERFAFNLLYRTPRYSN                  398
Cocoa vicilin:    YSNQNGRFFEACPEDFSQFQNMDVAVSAFKLNQGAIFVPHYNSKATF                  405

Cotton vicilin:   QNGRFYEACPREFRQLSDINVTVSALQLNQGSIFVPHYNSKATFVVL*                 445
Cocoa vicilin:    VVFVTDGYGYAQMACPHLSRQSQGSQSGRQDRREQEEESEEETFGEF                  452

Cotton vicilin:   VNEGNGYVEMVSPHLPRQSSFEEEEEQQQEQEQEEERRSGQYRKIRS                  492

Cotton vicilin:   QLSRGDI*                                                        499
```

Fig. 6

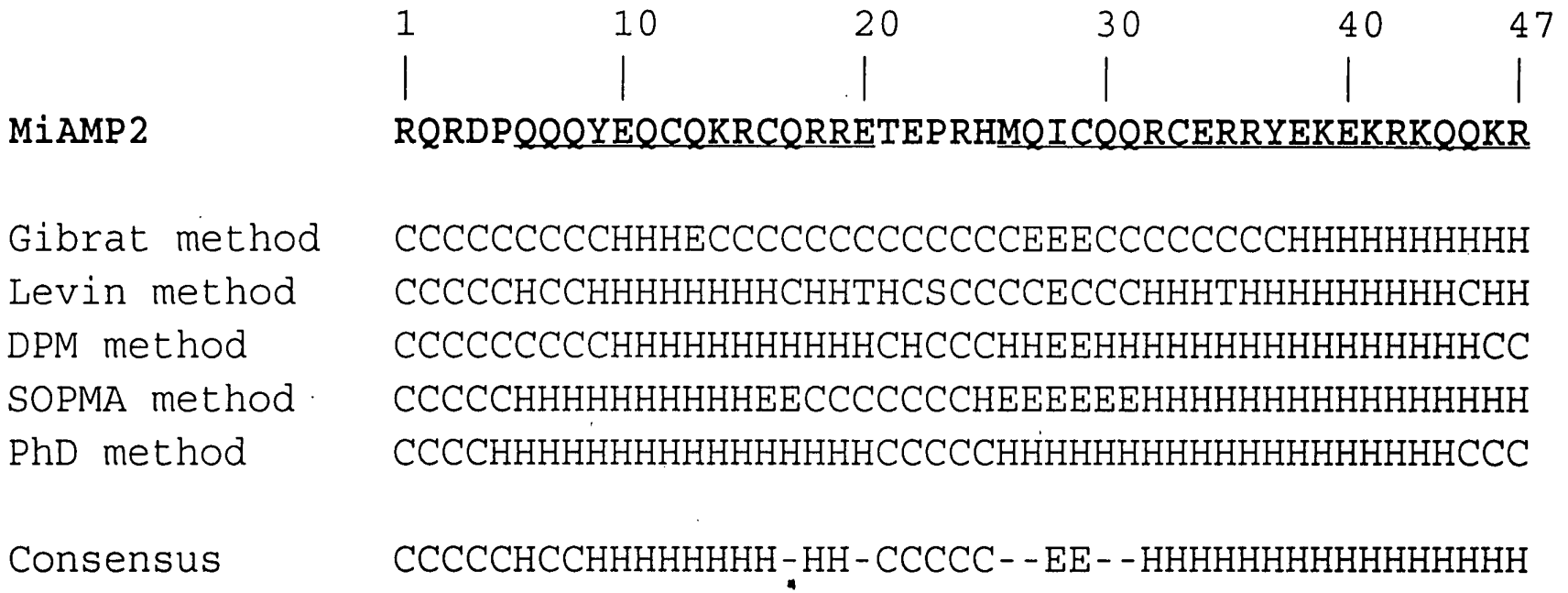
1 10 20 30 40 47
MiAMP2 RQRDPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQKR
Gibrat method CCCCCCCCCHHHECCCCCCCCCCCCCEEECCCCCCCCHHHHHHHHHH
Levin method CCCCCHCCHHHHHHHHCHHTHCSCCCCECCCHHHTHHHHHHHHHCHH
DPM method CCCCCCCCCHHHHHHHHHHHCHCCCHHEEHHHHHHHHHHHHHHHHCC
SOPMA method CCCCCHHHHHHHHHHEECCCCCCCHEEEEEEHHHHHHHHHHHHHHHHH
PhD method CCCCHHHHHHHHHHHHHHHHCCCCCHHHHHHHHHHHHHHHHHHHCCC
Consensus CCCCCHCCHHHHHHHH-HH-CCCCC--EE--HHHHHHHHHHHHHHHH

Fig. 7

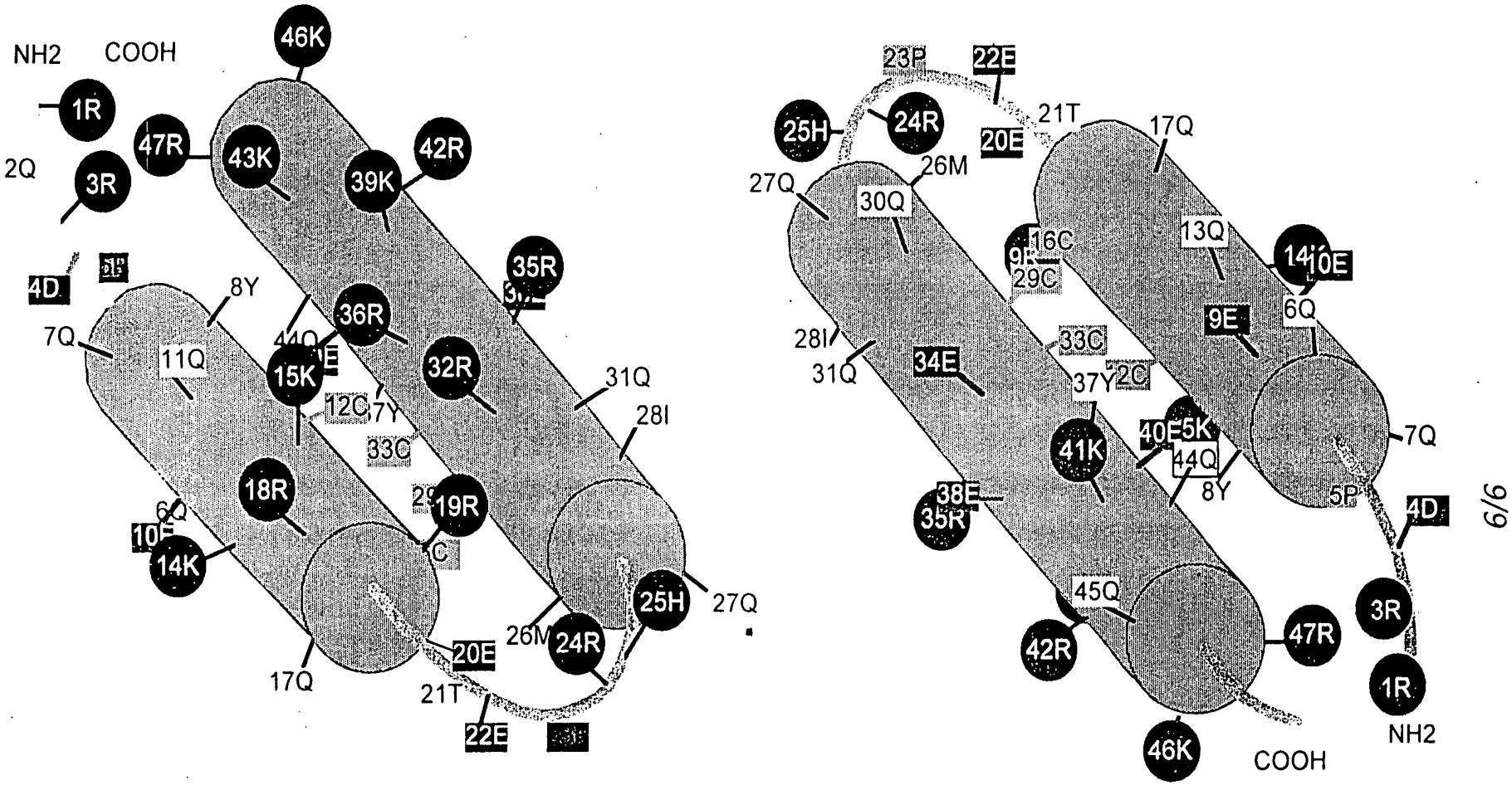
NH2
COOH
1R
2Q
3R
47R
43K
46K
39K
42R
4D
8Y
7Q
11Q
15K
36R
32R
35R
31Q
28I
12C
7Y
33C
18R
19R
14K
25H
27Q
26M
24R
20E
17Q
21T
22E
23P
22E
25H
24R
21T
20E
17Q
26M
27Q
30Q
16C
29C
13Q
10E
9E
6Q
28I
31Q
34E
33C
37Y
2C
41K
40E
5K
44Q
8Y
7Q
38E
35R
5P
4D
45Q
42R
47R
3R
1R
NH2
46K
COOH

Fig. 8